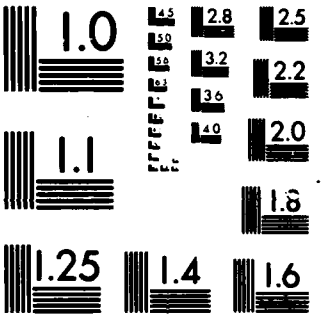
AD-A163 700 APPLICATION OF SOME NEW ARTIFICIAL NEURAL SYSTEM PROCESSING PRINCIPLES TO..(U) TRW RANCHO CARMEL ARTIFICIAL INTELLIGENCE CENTER SAN DIEGO CA. 1/1

UNCLASSIFIED R HECHT-NIELSEN ET AL. 18 NOV 85 F/G 6/4 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

MASTER COPY - FOR REPRODUCTION PURPOSES

(2)

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ARO 22323.1-MA	2. GOVT ACCESSION NO. N/A	3. RECIPIENT'S CATALOG NUMBER N/A
4. TITLE (and Subtitle) Application of some new artificial neural system processing principles to pattern classification		5. TYPE OF REPORT & PERIOD COVERED interim technical report 8/26/85 thru 11/18/85 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Robert Hecht-Nielsen Carolyn Smith		8. CONTRACT OR GRANT NUMBER(s) DAAG29-85-C-0025
9. PERFORMING ORGANIZATION NAME AND ADDRESS TRW Rancho Carmel AI Center One rancho Carmel San Diego, CA 92128		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE 18 November 85 13. NUMBER OF PAGES 11
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

DTIC ELECTE FEB 07 1986 S D

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Artificial neural systems
Pattern recognition
Artificial Intelligence
Adaptive networks

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The overall goal of the ARO study is to provide a firm conceptual and theoretical foundation for future work in ANS by exploring the use of three information processing principles. This entails both experimental exploration and theoretical development of "design laws" that can quantitatively predict system behavior and performance. This interim report describes the pattern environment which will support the planned experimentation. The pattern environment is a simple version of spoken english (called Lenglish), (con't)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-A163 780

DTIC FILE COPY

86 2 7 150

where each letter is associated with an arbitrary "sound" spectrum and there is a continuous transition between the spectra of adjacent letters. The signal input to the ANS will be the Lenglish reading of selected source texts.

The planned experimentation will use a model of unsupervised learning to discover feature detectors for spatiotemporal segments of the input language.

ARO 22323.1-MA

Application Of Some New Artificial Neural System Information Processing Principles To Pattern Classification

Interim Technical Report

Robert Hecht-Nielsen
Carolyn Smith

18 November 1985

US Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709

Contract No. DAA29-85-C-0025

TRW Rancho Carmel AI Center
One Rancho Carmel
San Diego, CA 92128
619-592-3206
ARPANET: HNIELSEN @ISI

Approved For Public Release;
Distribution Unlimited

1 Introduction

This report documents the progress to date on the Army Research Office/TRW *Application Of Some New Artificial Neural System Information Processing Principles To Pattern Classification* study. An important milestone has been reached in that we have selected the pattern environment and have developed a pre-processing tool which generates that environment for experimentation.

1.1 ARO Study Overview

The overall goal of the ARO study is to provide a firm conceptual and theoretical foundation for future work in ANS. In this effort we are exploring the use of three information processing principles (hierarchical template pattern storage, temporal compression, and simultaneous competitive template matching) and developing concise mathematical formulations of these principles. This will be accomplished with 6 tasks:

- The primary purpose of Task 1 is to select the pattern environment which supports the experimentation and develop the necessary pre-processing. This part of Task 1 has been completed and is documented in this interim report. Additional work on pre-processing techniques such as the Fourier-Mellin transform will be done as the study progresses.
- The work under Task 2 will use the signal environment to carry out experiments with a multi-slab ANS.
- In Task 3, data reduction and analysis will be conducted using the experimental data generated in Task 2.
- In Task 4, theoretical principles of ANS information processing will be developed, using the analysis of Task 3.
- Tasks 5 and 6 will docmument the work done in final and interm reports.

The schedule for these tasks is given in Figure 1.

1.2 Signal Environment For Experimental Work

The pattern environment which will be the substrate for the experimental work has been selected as a result of our work under task 1. We have decided to use a simple version of spoken English, which we call Letter-English or Lenglish. A description of Lenglish is given in section 2. The initial signal input to the experimentation will be the Lenglish reading of selected source texts, described in section 3. Section 4 of this report outlines our experimental plans for exploring the three processing principles using this pattern environment.

QUALITY INSPECTED 8

Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

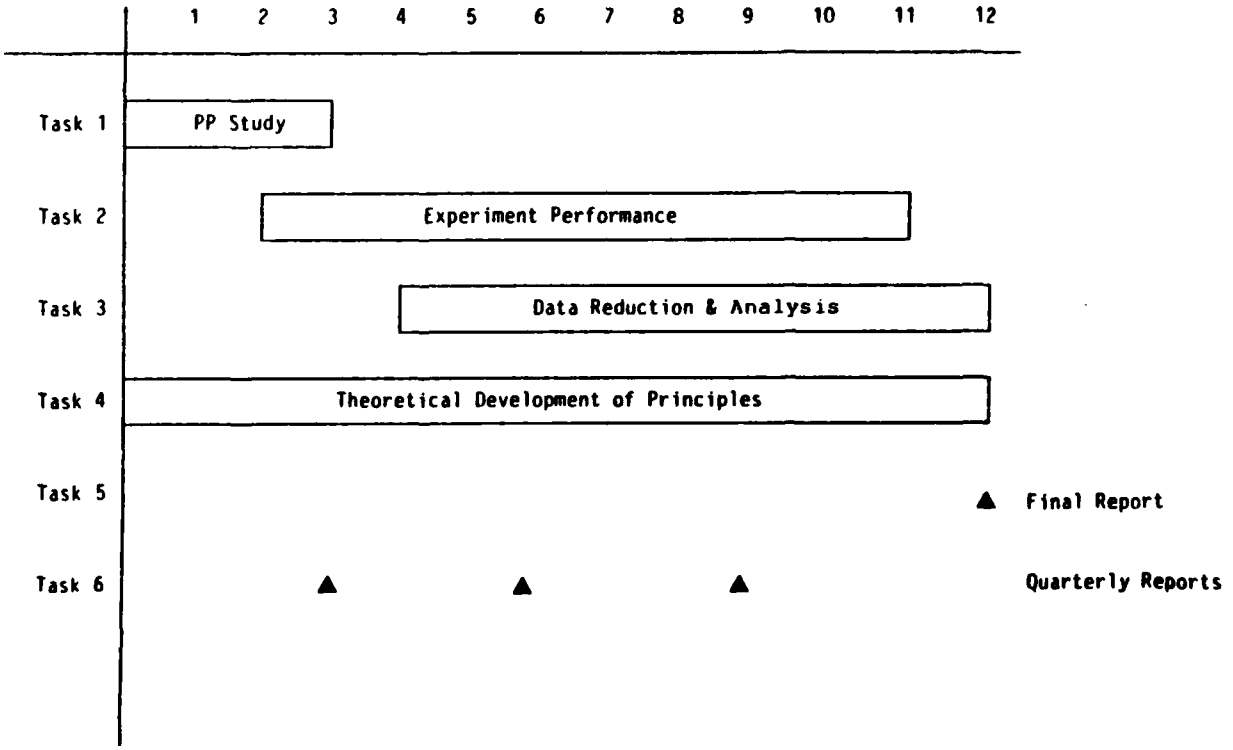


Figure 1: Schedule

2 Lenglish Description

The basic concept of Lenglish is to associate an arbitrary "sound" spectrum with each letter in the English language. Specifically, each letter in the alphabet is assigned one vector from a set of nearly orthogonal 32-dimensional vectors. The zero vector is used to represent pauses in speech between words. Pauses can be removed, if desired, to simulate the effect of word merging. To simulate the continuous nature of spoken language there is a smooth transition between adjacent letters within a word. The transition vectors are generated by convex combination of the adjacent letters' vectors. All of the vectors are normalized to "unit" length except the zero vector and transition vectors to or from the zero vector.

A pre-processing tool generates the signal input to the ANS by translating English text into Lenglish. The tool first finds the spectrum associated with a letter in the text stream. This spectrum is normalized and then repeated or "sampled" a user specified number of times. The tool then generates a user specified number of transition vectors to the spectrum of the next letter in the text. Blanks in the text seperating words represent pauses in speech and so are translated into the zero vector.

A movie has been made which shows the input signal generated by the pre-processing tool. Figure 2 shows one frame from the movie. The text rolls accross the screen from right to left. A fixed arrow indicates the current letter or transition being processed. The window below the text displays the spectrum for the current letter or transition.

3 Text Sources

In selecting the initial source text to be translated into Lenglish, we wanted a source with a restricted vocabulary, relatively short words, and a lot of repetition. For this reason we decided to use the text of childrens books to define the language. We have chosen the classic "The Little Engine That Could" by Watty Piper and the popular "Frog And Toad Are Friends" by Arnold Lobel. "The Little Engine That Could" contains 1349 words and uses a vocabulary of 302 words. The letter frequencies are given in Table 1. Most of the words are used infrequently, however about 10 percent of the words (24 words) are repeated at least 10 times in the text. The vocabulary set, with associated frequencies, is shown in Table 2. The second book, "Frog And Toad Are Friends", contains 2702 words and uses a vocabularly of 412 words. Over 20 of the words are repeated 20 times or more in the book. The letter and word frequencies are given in Tables 1 and 3.

4 Experimental Plans

The Lengish signal input derived from the source texts will be divided into training sets and test sets. These sets are the input stimulus to a multi-slab ANS such as the design sketched in figure 3. Our approach to the ANS experimentation with the pattern environment uses a model of unsupervized learning with competition. In this model, a competitive mechanism will be used to discover a set of feature detectors for the input stimulus. This is essentially the competitive learning of Rumelhart [1] with an important extension. Rumelhart's work demonstrated how an

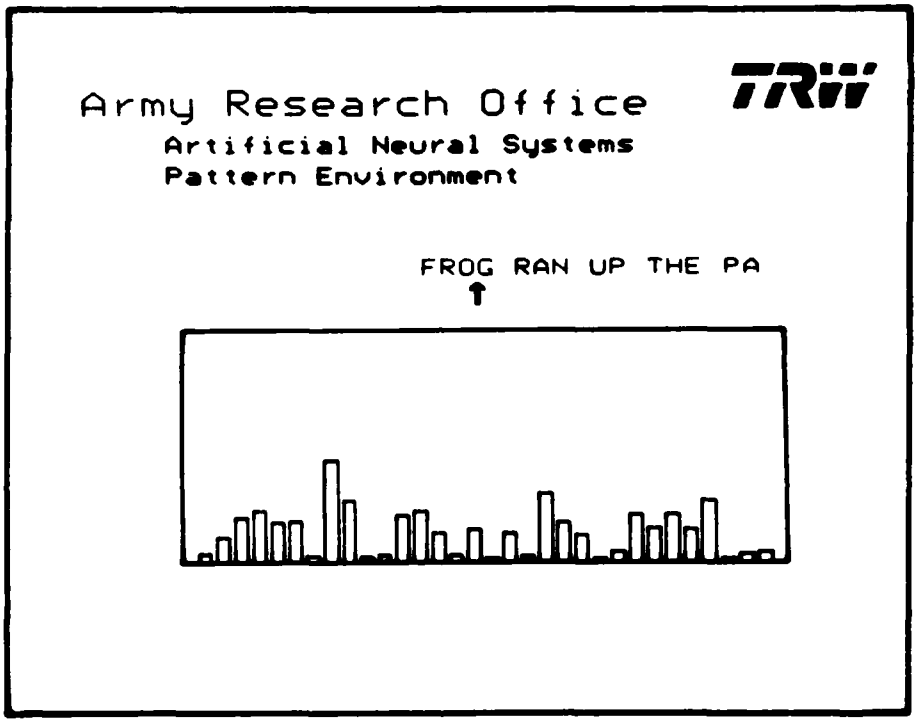


Figure 2: Input to ANS

Letter	Freq in "Engine"	Freq in "Frog"
A	309	772
B	67	136
C	97	109
D	255	484
E	577	885
F	85	229
G	138	234
H	285	496
I	318	502
J	7	9
K	40	116
L	312	328
M	62	170
N	384	504
O	380	880
P	94	115
Q	1	3
R	221	408
S	265	442
T	447	915
U	165	262
V	39	59
W	100	222
X	0	6
Y	125	182
Z	2	4
.	86	162
.	97	282

Table 1: Frequency Of Letter Use

ANS can organize itself into feature detectors for spatial patterns. Our work under this study will extend these results to the case of time varying (spatio-temporal) patterns. During experimentaion the time-varying signal of the simulated speech will be presented to the ANS. When the system reaches equilibrium, slab 1 will have discovered feature detectors for pattern primatives, possibily letters. The output of the slab 1 feature detectors are then the input to slab 2 which uses the same type of mechanism to discover feature detectors for higher level patterns, possibly letter pairs or word fragments. Similarly, each slab in the multi-slab system uses the same basic design to discover features in the input from the lower level slab. Success in self organization is achieved if the system reaches equilibrum. The training sets are used to allow the system to self-organize and the test sets are used to determine if the system has indeed equilibrated on features of the Lenglish source.

Note that this application will simultaneously use all three of the information processing principles of interest to this study. Hierachial pattern storage will be used to code primative features of the pattern (i.e. letters) in the lowest level, and composite features such as letter sequences at higher levels in the hierarchy. Temporal compression will be used to reduce the time variablity of higher level patterns. The recognition of time-varying patterns will occur as the avalanches compete for the privilege of being feature detectors for spatiotemporal segments of the input language.

Results of these experiments, as well as our theoretical results aimed at deriving "design laws" that can quantitatively predict system behavior and performance, will be given in later Interim Technical Reports and in the Final Report.

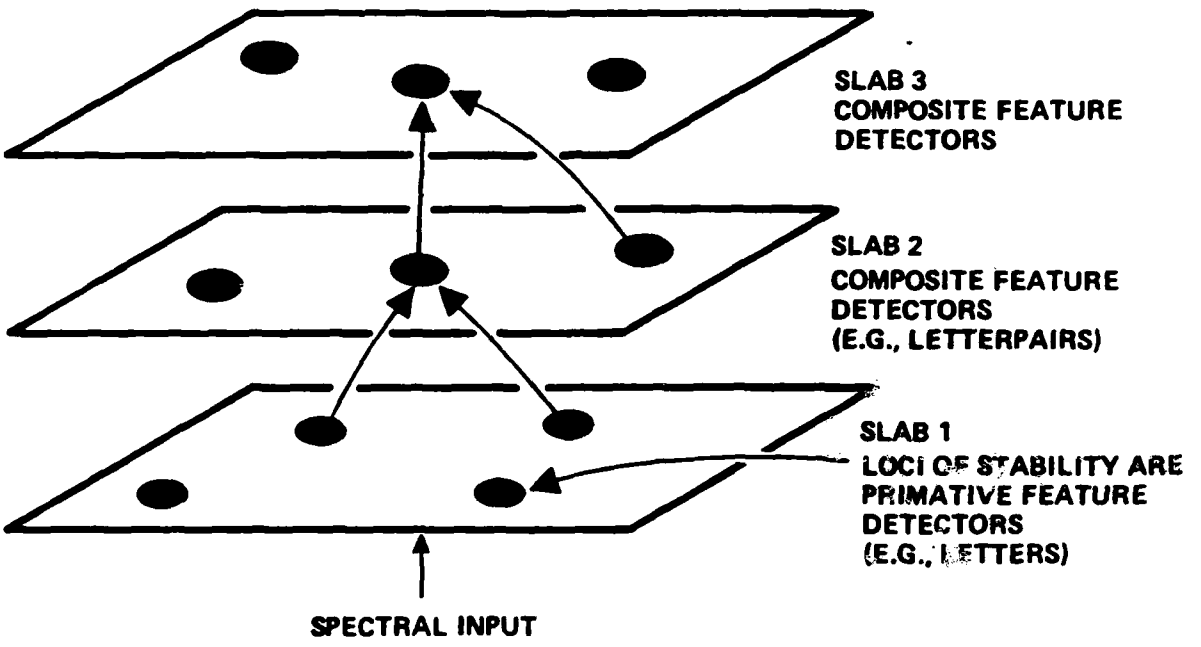


Figure 3: Multi-Slab ANS

, (comma)	86	CITY	2	HELPED	2	ON	7	STOPPED	3	WONT	8
. (period)	97	CLIMED	1	HER	2	ONE	4	STRONG	2	WORLD	4
A	16	CLOWN	9	HERE	4	ONLY	3	SUCH	1	YARD	1
ABOARD	1	CLOWNS	1	HERSELF	1	OR	6	SUDDEN	1	YELLOW	1
AFTER	1	COMES	2	HIM	2	ORANGES	1	SWITCHING	1	YOU	14
AIRPLANES	1	COMFORTABLE	1	HIS	2	OTHER	7	TEARS	1	YOURS	1
ALL	14	COMING	2	HITCHED	1	OUR	8	TEDDY	1		
ALMOST	1	COULD	8	HOW	1	OUT	4	THAN	1		
ALONG	2	CREAMY	1	HUNGRY	1	OVER	9	THAT	1		
AM	4	CRIED	6	HURRAY	2	PARLOR	1	THE	93		
AND	60	CRY	1	I	52	PASSENGER	2	THEIR	2		
ANIMALS	2	CURLS	1	IM	1	PEOPLE	2	THEN	5		
ANOTHER	4	DING	2	IMPORTANT	1	PEPPERMINT	1	THERE	3		
ANY	5	DINGY	1	IN	8	PERHAPS	1	THESE	2		
APPLES	1	DINNERS	1	INCH	1	PICTURE	1	THEY	4		
ARE	1	DINNING	1	INDEED	3	PLATE	1	THING	1		
ARM	1	DO	1	INDIGNANTLY	1	PLAY	5	THINGS	3		
AS	2	DOLLS	14	IS	6	PLEASE	9	THINK	11		
ASK	2	DONG	2	JACK	1	PRINT	1	THOSE	1		
ASKED	1	DOWN	6	JERK	1	PUFF	5	THOUGHT	7		
AT	2	DREAMED	1	JOLLY	1	PUFFED	3	TIRED	2		
AWAKE	1	DROPS	1	JUMPED	2	PULL	8	TO	27		
BABY	1	EAT	7	JUST	2	PULLED	3	TOGETHER	3		
BE	1	ELEPHANT	1	KIND	3	PUZZELS	1	TOP	1		
BEARS	1	ENGINE	33	KINDLY	1	QUICKLY	1	TOPS	1		
BECAUSE	1	ENGINES	2	KNIVES	1	REACHED	1	TOY	9		
BEEN	1	EVEN	2	LAST	1	READ	1	TOYS	15		
BEFORE	1	EVER	2	LAY	1	READY	1	TRACKS	1		
BEGAN	1	EVERY	1	LET	2	RED	1	TRAIN	16		
BELLOWED	1	EYES	3	LIKE	2	REST	1	TREATS	1		
BERTHS	1	FASTER	4	LITTLE	33	ROUNDHOUSE	3	TRIED	2		
BIG	10	FELT	1	LIVE	1	RUMBELD	1	TUGGED	2		
BLUE	8	FILLED	2	LOAD	1	RUMBELED	1	TURN	1		
BOBBED	1	FINE	1	LOADED	1	RUSTY	2	UNLESS	5		
BOOKS	2	FLAG	3	LOLLYPOPS	1	SAD	3	UNTIL	1		
BOTTLES	1	FOOD	6	LONG	1	SAID	4	UP	6		
BOYS	11	FOR	8	LOOK	1	SAW	3	US	13		
BREAKFASTS	1	FREIGHT	3	LOOKED	1	SAY	1	USE	1		
BRING	1	FRESH	1	LOOKS	1	SEEMED	1	VALLEY	1		
BROKEN	4	FRIENDS	2	MACHINES	2	SHE	17	VERY	9		
BROWN	2	FULL	2	MATTER	1	SHINY	3	WAITERS	1		
BUSY	1	FUNNIEST	1	MAYBE	1	SIDE	7	WANT	2		
BUT	8	FUNNY	2	ME	1	SIGHED	1	WAS	4		
CALLED	2	GET	1	MEAL	1	SIMPLY	1	WAVED	2		
CAME	2	GIRAFFES	1	MERRILY	2	SIT	1	WE	1		
CAN	19	GIRLS	11	MILK	1	SLEEPING	1	WEARY	1		
CANNOT	1	GLASS	1	MORE	1	SLOWLY	3	WENT	1		
CAR	1	GO	1	MOUNTAIN	14	SMILE	1	WERE	8		
CARRIED	1	GOING	1	MUST	2	SMILED	1	WHAT	2		
CARRY	1	GOLDEN	1	MY	3	SNORTED	1	WHATEVER	1		
CARRYING	1	GOOD	14	NECKS	2	SO	5	WHEELS	2		
CARS	6	GREAT	1	NEVER	1	SOFT	1	WHEN	2		
CHAIRS	1	GROWN	1	NEW	3	SOME	1	WHERE	2		
CHEEKED	1	HAD	2	NEWSPAPERS	1	SORTS	1	WHICH	1		
CHEER	2	HAPPY	2	NO	1	SPINACH	1	WHO	2		
CHILDREN	1	HAS	4	NOT	15	STARTED	1	WILL	3		
CHUG	6	HAVE	8	OF	16	STEADILY	1	WINDOWS	1		
CHUGGING	2	HE	4	OFF	4	STEAMED	1	WITH	14		
		HEADS	1	OH	2	STONG	1	WITHOUT	1		
		HELP	9	OLD	3	STOP	1	WONDERFUL	2		

Table 2: Word Frequency for "Engine"

,	155	CALENDAR	3	FLOOR	1	KNOCKED	1	ON	22	SITTING	2	TOGETHER	3		
.	278	CALLED	1	FOG	1	KNOW	2	ONCE	1	SKIP	1	TOLD	1		
A	61	CAME	6	FOR	21	LARGE	2	ONE	11	SLAMMED	1	TOO	4		
ABOUT	1	CAN	3	FORG	1	LAST	1	ONLY	1	SLEEP	1	TOOK	2		
ACROSS	1	CANNOT	1	FOUND	4	LATER	1	ONTO	3	SLOWLY	1	TOP	3		
AFTER	1	CATCHING	1	FOUR	2	LAUGHED	5	OUT	20	SMALL	3	TORE	3		
AFTERNOON	1	CLEAR	1	FRIEND	3	LAUGHING	2	OUTSIDE	1	SNAIL	7	TREE	1		
AGAIN	2	CLIMED	2	FRIENDS	1	LET	2	PAGE	9	SNAKE	4	TROUBLE	1		
AGAINST	4	CLOSED	2	FROG	88	LETTER	12	PAPER	3	SNEEZE	1	TURTLE	7		
ALL	9	CLOTHES	1	FROM	6	LIGHT	1	PAST	1	SNOW	1	TWO	4		
ALONG	3	COLD	1	FRONT	6	LIKE	1	PATH	1	SO	3	UNHAPPY	1		
ALWAYS	2	COLDER	2	FUN	1	LISTEN	1	PATHS	1	SOME	2	UNTIL	3		
AM	11	COME	4	FUNNY	6	LITTLE	1	PEEK	1	SOMEONE	2	UP	12		
AN	1	CORNER	1	GAVE	2	LIZARDS	3	PENCIL	1	SOMETHING	1	UPON	1		
AND	55	COULD	9	GET	6	LONELY	1	PERHAPS	1	SOON	2	VERY	6		
ANGRY	1	COUNT	2	GETTING	1	LONG	6	PICKED	1	SORRY	1	VOICE	3		
ANSWER	3	COURSE	1	GLAD	1	LOOK	9	PIECE	1	SPARROW	2	WAILED	1		
ANY	2	COVERED	1	GLASS	1	LOOKED	7	PLEASE	3	SPLASHES	2	WAIT	3		
ANYMORE	1	COVERS	2	GLASSES	1	LOOKING	5	PLEASED	1	SPRING	2	WAITED	2		
ANYONE	1	CRAWLED	1	GO	8	LOOKS	2	POCKET	6	SQUARE	2	WAITING	3		
ANYTHING	2	CRIED	5	GOING	1	LOSE	1	PORCH	8	STAND	1	WAKE	4		
APRIL	3	CUP	1	GOOD	3	LOST	1	POUR	1	STANDING	1	WALK	2		
ARE	7	DARK	2	GOT	3	LOT	1	POURED	3	STARS	1	WALKED	7		
ARTE	1	DAY	5	GRASS	2	LYING	1	POURING	1	STAY	1	WALL	4		
AS	2	DAYS	1	GREEN	3	MADE	4	PULLED	2	STILL	3	WANT	4		
ASKED	10	DEAR	1	HAD	5	MAIL	7	PUSHED	2	STOOD	3	WANTS	1		
ASLEEP	4	DECEMBER	1	HALF	1	MAILBOX	3	PUT	10	STORY	19	WARM	1		
AT	9	DID	5	HAPPY	1	MAKES	2	QUITE	1	SUIT	11	WAS	26		
AWAY	7	DO	9	HARD	1	MANY	1	RACCOON	1	SUMMER	1	WATER	10		
BACK	10	DOES	2	HAS	3	MARCH	1	RAN	6	SUN	2	WE	5		
BANG	2	DONE	1	HAVE	11	MATTER	1	REALLY	1	SURE	1	WEAR	1		
BANGED	1	DONT	5	HEAD	10	MAY	7	REST	1	SWAM	5	WEARING	1		
BANGING	1	DOOR	2	HEARD	1	ME	20	RESTING	1	SWIM	1	WELL	6		
BATHING	11	DOWN	8	HELP	5	MEADOW	2	RIGHT	4	TAKE	1	WENT	11		
BE	6	DRAGONFLIES	2	HERE	8	MEADOWS	1	RIVER	6	TAKING	1	WERE	4		
BEAUTIFUL	1	DRANK	1	HIM	8	MEANS	1	RIVERBANK	2	TALL	1	WHAT	7		
BECAUSE	7	DRAT	1	HIS	32	MELTING	1	ROCKS	2	TEA	2	WHEN	4		
BED	10	DRIPPED	1	HOLED	1	MINE	1	ROOM	1	THEY	1	WHERE	1		
BEEN	2	EMPTY	1	HOLES	3	MISS	1	ROUND	1	TELL	9	WHILE	1		
BEFORE	1	END	1	HOME	5	MORE	2	RUN	1	TERRIBLE	1	WHITE	2		
BEGAN	2	ENOUGH,	1	HOPE	3	MOUSE	3	SAD	4	THAT	28	WHOLE	2		
BEGIN	1	ENVELOPE	2	HOT	1	MUCH	1	SAID	75	THE	127	WHY	10		
BEGINNING	1	EVEN	1	HOUSE	11	MUD	1	SAT	3	THEM	3	WILL	21		
BEHIND	3	EVENINGS	1	HOW	4	MUST	2	SAW	2	THEN	16	WIM	1		
BEST	2	EVER	4	HURRIED	1	MY	23	SCREAMED	1	THERE	10	WINDOW	4		
BETTER	2	EVERY	1	HURT	2	NAP	1	SEE	8	THESE	1	WITH	2		
BIG	3	EVERYONE	1	I	52	NEARBY	1	SEEN	1	THEY	12	WOODS	3		
BLACK	2	EXCUSE	1	IF	6	NEED	1	SEND	4	THICK	2	WORLD	2		
BLAH	1	EYES	1	IN	25	NEVER	3	SENT	3	THIN	2	WORRY	1		
BLINKED	1	FALLEN	2	INSIDE	1	NEW	1	SEWED	2	THINK	18	WOULD	1		
BOX	1	FAST	1	INTO	8	NEXT	1	SEWING	1	THINKS	1	WRITE	1		
BRIGHT	1	FEBRUARY	1	IS	30	NIGHT	1	SHELF	1	THIS	3	WROTE	3		
BUT	16	FEEL	2	IT	18	NO	9	SHINING	1	THOUGHT	3	YEAR	1		
BUTTON	29	FEELING	6	JACKET	3	NONE	1	SHIVER	1	THROUGH	2	YES	5		
BUTTONS	5	FEET	2	JANUARY	1	NOT	35	SHOULD	2	TIME	9	YET	2		
		FELL	1	JOY	1	NOVEMBER	3	SHOUTED	2	TIRED	2	YOU	34		
		FIELD	2	JUMPED	3	NOW	5	SHUTTERS	1	TO	44	YOUR	10		
		FILED	2	JUST	1	OF	44	SILLY	2	TOAD	112				
		FIND	1	KEEP	1	OFF	5	SINCE	1	TOADS	8				
		FLEW	1	KNEW	1	OH	4	SIT	1	TODAY	3				

Table 3: Word Frequency for "Frog"

References

[1] Rumelhart, D. E. and D. Zipser, " "Feature Discovery by Competitive Learning" *ICS Report No. 8407*. June 1984.

END

FILMED

3-86

DTIC